Game Theory
MOT Seminar

John Musacchio
johnm@soe.ucsc.edu
Who am I?

- John Musacchio
- Assistant Professor in ISTM
  - Joined January 2005
- PhD from Berkeley in Electrical Engineering
- Research Interests
  - Game Theory applied to pricing problems.
  - Modeling and control of networks
- Experience
  - 2$\frac{1}{2}$ years at a Start-Up
    - Helped design a chip-set for computer-networking switches
What is Game Theory?

- Study of interacting strategic agents.
- Used frequently in economics and other sciences.
  - Competition between firms.
  - Auction Design.
  - International Policy.
  - Evolution of Species.
  - And many more...
Example

- How should they conduct their transaction?
  - Pre-pay?  ➔ Access Point might cheat.
  - Post-pay?  ➔ Client might cheat.
  - Pay as she goes?

- Will this payment model work?
  - Will access point (AP) change its price over session duration?
  - Will client and AP accept this model?
Motivation

- Many public access points, but far from universal coverage.
- Enormous growth in private access points.
- Incentivize AP owners to open up to public. → Universal Coverage.
Classic Example: Prisoner's Dilemma

Prisoner A

Prisoner B

<table>
<thead>
<tr>
<th></th>
<th>Silent</th>
<th>Betray</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silent</td>
<td>(-1, -1)</td>
<td>(-4, 0)</td>
</tr>
<tr>
<td>Betray</td>
<td>(0, -4)</td>
<td>(-3, -3)</td>
</tr>
</tbody>
</table>
Competing Firms

<table>
<thead>
<tr>
<th></th>
<th>Stagnate</th>
<th>Innovate</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Firm A</strong></td>
<td>(2, -1)</td>
<td>(3, 0)</td>
</tr>
<tr>
<td><strong>Firm B</strong></td>
<td>(-1, 3)</td>
<td>(0, 0)</td>
</tr>
</tbody>
</table>
**Elements of a Game**

- **Players**
- **Strategy**
  - A Player’s action
    - Innovate or Stagnate
  - Strategy Space - Set of all possible actions
  - Strategy Profile - Particular combination of player strategies.
- **Payoff**
  - A mapping from player strategy profile to player rewards
    - Example: \( U( (I, I) ) = (0,0) \)
Solution Concept

Nash Equilibrium

- A strategy profile from which no player has an incentive to deviate unilaterally

- Example \((I, I)\) is a NE

  - \(U_A(I, I) > U_A(S, I)\) Firm A cannot do better by deviating
  - \(U_B(I, I) > U_B(I, S)\) Firm B cannot do better by deviating

- \((I, I)\) is a Nash Equilibrium.

Do all games have a Nash Equilibrium?
Example: Leader, Imitator (matching pennies)

- Idea: Player 1 (Imitator) wants a match, Player 2 (Leader) doesn’t.
- What is the Nash Equilibrium?
- Expand strategy space to allow randomized or “mixed” strategies.

<table>
<thead>
<tr>
<th></th>
<th>Product A</th>
<th>Product B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imitator Deviates</td>
<td>(1,-1)</td>
<td>(-1, 1)</td>
</tr>
<tr>
<td>Leader Deviates</td>
<td>(-1,1)</td>
<td>(1,-1)</td>
</tr>
</tbody>
</table>

Leader

Product A

Product B
Example: Leader, Imitator

- NE is not strict in this case.
  - (At NE, players are indifferent to switching)
  - Such an NE is said to be not strict
Nash Existence

- Finite Strategy Space (J.F. Nash 1950)
  - Every n-player game has at least one Nash Equilibrium (possibly mixed).
Static vs. Multi Stage

- **Static Games**
  - Players choose strategies simultaneously, without knowing what the others do.

- **Multi-Stage**
  - Game is played in multiple rounds.
  - Players may see how others played in previous rounds.
    - That information helps choose how to play in the next round.
  - A strategy is a full specification of what actions to take in each stage, as a function of the observations from previous stages.
Repeated Innovating Firms Game
(Repeated Prisoners Dilemma)

- **Recall, Both firms innovate in the one-shot game.**
  - Combined reward is 0.
  - If they had both stagnated instead, their combined reward would have been 4.

- **What happens if the game is repeated?**
  - Same game is repeated every year forever.
  - NPV of future payoffs is discounted by a discount factor $\beta$. 
Aside: What is Net Present Value?

How much is $1 worth to you 1 year from now?

Something less than a $1 ... say $β.$
Repeated Innovating Firms Game
(Repeated Prisoners Dilemma)

- Firm A’s:
  \[ \sum_{n=0}^{\infty} \beta^n A(x_n, y_n) \]

- Firm B’s:
  \[ \sum_{n=0}^{\infty} \beta^n B(x_n, y_n) \]

- Where
  \( x_n \) = Albert’s action in slot \( n \)
  \( A(\ ) \) = Albert’s payoff function
  \( y_n \) = Bob’s action in slot \( n \)
  \( B(\ ) \) = Bob’s payoff function
  \( \beta \) = Discount factor
Repeated Innovating Firms Game

- **Strategy:** I will stagnate as long as you do.
- **Threat:** If you choose to innovate once, I will innovate forever thereafter.

- This is a NE if $\beta > 1/3$
Repeated Innovating Firms Game

Proof:
- Suppose at time $t$, A innovates
  - B retaliates by innovating forever thereafter
  - A is forced to innovate at times $t+1$, $t+2$, ... as well
- A’s net payoff

\[
\Delta = \beta^t (3 - 2) + \sum_{n=t+1}^{\infty} \beta^n (0 - 2)
\]

\[
= \beta^t - 2 \frac{\beta^{t+1}}{1 - \beta}
\]

\[
= \frac{\beta^t}{1 - \beta} (1 - \beta - 2\beta) \leq 0 \text{ when } \beta \geq \frac{1}{3}
\]
Repeated Innovating Firms Game

- **Intuition**
  - When $\beta$ is large, future consequences of breaking the collusion agreement outweigh short term gain.
  - When $\beta$ is small, short term gain is more important than long term consequences.
Player 2 can “threaten” to choose R in stage 2 to get Player 1 to pick R is stage 1.
- But in the subgame starting in slot 2, Player 2 is compelled to pick L.
- Player 2’s threat is not credible.
- \((R,R)\) is indeed a NE, but not SPE.

Only \((L,L)\) is a SPE.
We said that the following strategy profile is a Nash Equilibrium:

- **Strategy:** I will stagnate as long as you do.
- **Threat:** If you choose to innovate once, I will innovate forever thereafter.

**Is it a SPE?**

- Yes.
- In the subgame after the first deviation, it is rational to Innovate forever thereafter if you expect your opponent to do the same.
Repeated Innovating Firms Game

- “Folk Theorem” - Cooperating can be rational if games are repeated forever.
- However, threat strategies can be used to enforce other outcomes.

Claim: Any Reward vector in the green region can be enforced by an SPE.
Repeated Innovating Firms Game

Proof:

- Consider $v$ in the green region:
  \[ v = \sum_{j=1}^{4} \lambda_j r_j \quad v \geq (0, 0) \]

- Pick integers $N_j$ that satisfy
  \[ \lambda_j \approx \frac{N_j}{N} \text{ for } j = 1, \ldots, 4 \]
  \[ N = N_1 + \cdots + N_4 \]

- They agree to play
  - (I,I) the first $N_1$ steps
  - (S,S) the next $N_2$ steps, etc...
  - When someone deviates from the schedule, the other retaliates by playing I forever thereafter.

One shot rewards.
Finitely Repeated Innovating Firms Game

- Suppose they play the game only $N$ times.
  - Is it a SP Equilibrium to play $(S,S)$ in all turns?
- Consider the $N$th stage of the game.
  - In the $N$th stage, the players don’t have to worry about how their action affects the future.
  - Thus at the $N$th stage both players Innovate.
    - (This is a dominant strategy for both players.)

- At time $N-1$, the players know their actions can’t affect the future.
  - Thus $(I,I)$ is again the dominant strategy.
Finitely Repeated Innovating Firms Game

- By induction,
  - The players play (I,I) in every slot.
  
- Such a strategy profile is the only Sub-Game Perfect Nash Equilibrium (SPE).
Example: Cournot Competition

- Two firms produce goods
  - Choose quantities $q_1, q_2$
- Market clearing price
  - $A - q_1 - q_2$
- Cost of production is $C$ per unit
  - $U_1(q_1, q_2) = (A - q_1 - q_2)q_1 - Cq_1$
    - $= (A - C - q_1 - q_2)q_1$
    - $= (B - q_1 - q_2)q_1$
    - $B := A - C$
- Firm 1 Best Response
  - $q_1^* = \arg\max_{q_1} U_1(q_1, q_2) = \frac{B - q_2}{2}$
Cournot Competition

Firm 1: $q_1$

Firm 2: $q_2$

**NE:** \( \left( \frac{B}{3}, \frac{B}{3} \right) \)

**Payoffs:** \( \left( \frac{B^2}{9}, \frac{B^2}{9} \right) \)

**Firm 1 best response:** \( \frac{B - q_2}{2} \)

**Firm 2 best response:** \( \frac{B - q_1}{2} \)
Repeated Cournot Competition

**One Shot Result Summary:**
- Clearing Price is: \((A-q_1-q_2)\)
- Cost per unit: \(C\)
- \(B := A - C\)
- Equilibrium: \((B/3, B/3)\)
- Each Firm earns \(B^2/9\)

**But, Combined Payoff maximized when firms produce \(B/4\) goods each.**
- In this case each firm earns \(B^2/8\)
- An aside: why \(B/4\) is not played in the static game?
  - If firm \(A\) produces \(B/4\), it is more profitable for firm \(B\) to produce \(3B/8\) than \(B/4\)
  - Firm \(A\) then in turn produces \(5B/16\), and so on...
Repeated Cournot Competition

Q: Is it possible for two firms to reach an agreement to produce $B/4$ instead of $B/3$ each?

- **Consider the strategy:** I will produce $B/4$ as long as you do. If you deviate, I will produce $B/3$ forever thereafter.

A firm has two choices in each round:

- **Cooperate:** produce $B/4$ and make profit $B^2/8$
- **Cheat:** produce $3B/8$ and make profit $9B^2/64$

But in the subsequent rounds, cheating will cause
- its competitor to produce $B/3$ as punishment
- its own profit to drop back to $B^2/9$
Repeated Cournot Competition

Is there any incentive for a firm not to cheat?
Let’s look at the accumulated payoffs:

- If it cooperates:
  \[ S_c = (1 + \delta + \delta^2 + \delta^3 + \ldots) \frac{B^2}{8} = \frac{B^2}{8}(1 - \delta) \]

- If it cheats:
  \[ S_d = \frac{9B^2}{64} + (\delta + \delta^2 + \delta^3 + \ldots) \frac{B^2}{9} \]
  \[ = \left\{ \frac{9}{64} + \frac{\delta}{9}(1 - \delta) \right\} B^2 \]

So it will not cheat if \( S_c > S_d \). This happens only if \( \delta > \frac{9}{17} \).

- Conclusion
  - If future return is valuable enough to each player, then strategies exist for them to play socially efficient moves.
Wi-Fi Access Point Pricing as a Dynamic Game

John Musacchio and Jean Walrand
{musacchj,wlr}@eecs.berkeley.edu
University of California, Berkeley
Dissertation Talk
November 15, 2004
Outline

Wi-Fi Access Point Pricing as a Dynamic Game

- Motivation
- General Formulation
- Web Browsing Model
  - Multi-Hop Case
- File Transfer Model
  - Known Type
  - Unknown Type
- Bayesian Game
  - Bounded Length
  - Unbounded Length
- Conclusions

Achieving Fair Rates with Backpressure
Motivation

How should they conduct their transaction?
- Pre-pay? → Access Point might cheat.
- Post-pay? → Client might cheat.
- Pay as she goes?

Will this payment model work?
- Will access point (AP) change its price over session duration?
- Will client and AP accept this model?
Motivation

- Many public access points, but far from universal coverage.
- Enormous growth in private access points.
- Incentivize AP owners to open up to public.
  → Universal Coverage.

Motivation
Outline

Wi-Fi Access Point Pricing as a Dynamic Game

- Motivation
- General Formulation
  - Web Browsing Model
    - Multi-Hop Case
  - File Transfer Model
    - Known Type
    - Unknown Type
  - Bayesian Game
    - Bounded Length
    - Unbounded Length
- Conclusions
General Formulation

Discrete time slot model: \[ 1 \ 2 \ \ldots \ t \]

Access point proposes price at the start of a slot:

Client’s Choices:

Accept
Quit Game

Implicit Assumption: Client’s utility depends on continuity of connection.
Web Browsing Model

Client Temporal Utility

Slope: $U$

- Client knows the value of $(U, \tau)$.
- AP knows the distributions of $(U, \tau)$.
- Session Utility = $U \times \min(T, \tau)$, where $T = \#\text{slots client ends up buying}$. 

$T$: Intended Session Length
Web Browsing Model

AP Tradeoff:

- Charge too much, and risk the client leaving.
- Play it too safe, and not capture as much revenue as he could have.
Client Objective

- Client wants to finish session with positive net payoff.

\[(U - p_1) + (U - p_2) + \ldots + (U - p_{\min[T, \tau]})\]

- Intuitive “Myopic” Strategy:
  - Accept iff \( p_t \leq U \).
When a myopic client accepts a price, the AP learns a new lower-bound for $U$ (utility per slot).

- Should the AP’s price change from slot to slot?

Web Browsing - AP Learning
Suppose the Client follows a myopic strategy.

Then, the Access Point:
- Should never decrease price.
- Choose from non decreasing sequences \((P^+)\) to maximize:

\[
\max_{\{p_t\}_{1}^{\infty} \in P^+} \left[ \sum_{i=1}^{\infty} p_t P(U \geq p_t) P(\tau \geq t) \right]
\]

Sum Separates. Optimal Solution:
- \(p_t = p^*\) where \(p^*\) maximizes \(pP(U \geq p)\).
Nash Equilibrium

We have shown:

Client: Myopic

We observe that:

Access Point: Charge nondecreasing (fixed) price.

We have found a Nash Equilibrium.

Access Point: Charge nondecreasing (fixed) price.

Client: Myopic
Perfection?

First Slot Optimization:

Later Slot Optimization:

Web Browsing - Nash Equilibrium

John Musacchio
What about Perfection?

- Consider the continuation game beginning at slot $s$. AP wants to solve:

$$\max_{\{p_t\}_{t=s}^\infty} \left[ \sum_{t=s}^\infty p_t P(U > p_t \mid U > p_{s-1}) P(\tau \geq t \mid \tau \geq s) \right].$$

- Should not lower prices, thus Bayes’ rule yields

$$\max_{\{p_t\}_{t=s}^\infty} \left[ \sum_{t=s}^\infty p_t P(U > p_t) P(\tau \geq t) \right] \times \frac{1}{P(\tau \geq s) P(U > p_{s-1})}.$$

- Similar objective as before.
Web Browsing Model

Theorem 1

Assume:
- \( U, \tau \) independent, finite mean.

Then a Perfect Bayesian Equilibrium (PBE) is:
- AP charges \( p^* \in \arg\max_p pP(U \geq p) \).
- Client is myopic.
  - Myopic means: connect iff \( p_t \leq U \) and \( t \leq \tau \).
File Transfer Model

- Client’s utility a step function.

\[ F(T, \tau) = \begin{cases} 
0 & \text{if } T < \tau \\
U_\tau & \text{if } T = \tau 
\end{cases} \]
Suppose: AP Knows client type \((U, \tau)\).

- **Decision in last slot:**
  
  Net Payoff for entire game:

\[
U_\tau - p_\tau - p_{\tau-1} \cdots - p_1
\]

- **AP can charge up to** \(U_\tau\) **in the last slot!**
  - She should not be willing to pay anything until last slot!
  - *(A “Pessimistic” Strategy.)*
Outline

Wi-Fi Access Point Pricing as a Dynamic Game

- Motivation
- General Formulation
- Web Browsing Model
  - Multi-Hop Case
- File Transfer Model
  - Known Type
  - Unknown Type
- Bayesian Game
  - Bounded Length
  - Unbounded Length
- Conclusions
File Transfer Model: **Unknown** Type

- Client’s utility a step function.

\[
F(T, \tau) = \begin{cases} 
0 & \text{if } T < \tau \\
U_\tau & \text{if } T = \tau 
\end{cases}
\]

- **AP** does not know client type:
  - \(U\) distributed on \([l, h]\).
  - \(\tau\) distributed on \(\{1, \ldots, n\}\).
  - Client knows her sample values.
  - **AP** just knows the distributions.
    - \(U\) and \(\tau\) may not be independent.

File Transfer - Unknown Type

John Musacchio
Theorem

- In PBE:
  - Client: “Pessimistic:”
    - Before last slot: Don’t pay
    - At last slot: Willing to Pay
  - AP charges

\[
p_i = \begin{cases} 
0 & \text{if } t < t^* \\
u^*t^* & \text{otherwise.}
\end{cases}
\]

Where \((u^*, t^*)\) maximizes \(utP(U > u, \tau = t)\).
Proof Sketch

- $U$ has cont. dist., so look at $\epsilon$ sized regions of support.
- AP charges at least $n \times l$ in last slot
  - Clients of type $U \in [l, l + \frac{\epsilon}{n})$, $\tau = n$ are $\epsilon$ pessimistic: don’t pay more than $\epsilon$ in earlier slots.
- AP charges at least $nl + \epsilon$ to non-pessimistic clients.
  - Clients of type $U \in [l, l + \frac{2\epsilon}{n})$, $\tau = n$ are $\epsilon$ pessimistic
- Induction shows clients of all types are $\epsilon$ pessimistic.

\[ U, \text{Utility per slot} \]
\[ l \quad l + \frac{\epsilon}{n} \quad l + \frac{2\epsilon}{n} \quad \cdots \quad h \]
\[ T, \text{File Length} \]
\[ n \quad n-1 \quad \cdots \quad 1 \]

Client Type Space
Facing a pessimistic Client AP has one chance to charge.

- Guesses the best slot to try charging.

Inefficient

- If client finishes before the AP charges, she gets a free file.
- If the AP charges before the file is finished, the client leaves.
Final Remarks

- **Game Theory** important tool in conceptualizing strategic interactions
  - Between Competing Firms
  - Buyers and Sellers
  - Other Interacting Agents